

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: NI, et al.

(ii) TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR RECEPTOR
TR10

(iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
- (B) STREET: 9410 KEY WEST AVENUE
- (C) CITY: ROCKVILLE
- (D) STATE: MD
- (E) COUNTRY: US
- (F) ZIP: 20850

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: BROOKES, ANDERS A.
- (B) REGISTRATION NUMBER: 36,373
- (C) REFERENCE/DOCKET NUMBER: PF379PP

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (301) 309-8504
- (B) TELEFAX: (301) 309-8512

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3566 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 109..1266

60050976

(ix) FEATURE:

(A) NAME/KEY: sig_peptide

(B) LOCATION: 109..271

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: 274..1266

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGACCCACGC GTCCGCCAC GCGTCCGGAG AACCTTTGCA CGCGCACAAA CTACGGGGAC	60
GATTTCTGAT TGATTTTGG CGCTTCGAT CCACCCTCCT CCCTTCTC ATG GGA CTT	117
Met Gly Leu	
-55	
TGG GGA CAA AGC GTC CCG ACC GCC TCG AGC GCT CGA GCA GGG CGC TAT	165
Trp Gly Gln Ser Val Pro Thr Ala Ser Ser Ala Arg Ala Gly Arg Tyr	
-50 -45 -40	
CCA GGA GCC AGG ACA GCG TCG GGA ACC AGA CCA TGG CTC CTG GAC CCC	213
Pro Gly Ala Arg Thr Ala Ser Gly Thr Arg Pro Trp Leu Leu Asp Pro	
-35 -30 -25	
AAG ATC CTT AAG TTC GTC GTC TTC ATC GTC GCG GTT CTG CTG CCG GTC	261
Lys Ile Leu Lys Phe Val Val Phe Ile Val Ala Val Leu Leu Pro Val	
-20 -15 -10 -5	
CGG GTT GAC TCT GCC ACC ATC CCC CGG CAG GAC GAA GTT CCC CAG CAG	309
Arg Val Asp Ser Ala Thr Ile Pro Arg Gln Asp Glu Val Pro Gln Gln	
1 5 10	
ACA GTG GCC CCA CAG CAA CAG AGG CGC AGC CTC AAG GAG GAG GAG TGT	357
Thr Val Ala Pro Gln Gln Gln Arg Arg Ser Leu Lys Glu Glu Glu Cys	
15 20 25	
CCA GCA GGA TCT CAT AGA TCA GAA TAT ACT GGA GCC TGT AAC CCG TGC	405
Pro Ala Gly Ser His Arg Ser Glu Tyr Thr Gly Ala Cys Asn Pro Cys	
30 35 40	
ACA GAG GGT GTG GAT TAC ACC ATT GCT TCC AAC AAT TTG CCT TCT TGC	453
Thr Glu Gly Val Asp Tyr Thr Ile Ala Ser Asn Asn Leu Pro Ser Cys	
45 50 55 60	
CTG CTA TGT ACA GTT TGT AAA TCA GGT CAA ACA AAT AAA AGT TCC TGT	501
Leu Leu Cys Thr Val Cys Lys Ser Gly Gln Thr Asn Lys Ser Ser Cys	
65 70 75	
ACC ACG ACC AGA GAC ACC GTG TGT CAG TGT GAA AAA GGA AGC TTC CAG	549
Thr Thr Thr Arg Asp Thr Val Cys Gln Cys Glu Lys Gly Ser Phe Gln	
80 85 90	
GAT AAA AAC TCC CCT GAG ATG TGC CGG ACG TGT AGA ACA GGG TGT CCC	597
Asp Lys Asn Ser Pro Glu Met Cys Arg Thr Cys Arg Thr Gly Cys Pro	

95	100	105	
AGA GGG ATG GTC AAG GTC AGT AAT TGT ACG CCC CGG AGT GAC ATC AAG Arg Gly Met Val Lys Val Ser Asn Cys Thr Pro Arg Ser Asp Ile Lys 110 115 120			645
TGC AAA AAT GAA TCA GCT GCC AGT TCC ACT GGG AAA ACC CCA GCA GCG Cys Lys Asn Glu Ser Ala Ala Ser Ser Thr Gly Lys Thr Pro Ala Ala 125 130 135 140			693
GAG GAG ACA GTG ACC ACC ATC CTG GGG ATG CTT GCC TCT CCC TAT CAC Glu Glu Thr Val Thr Thr Ile Leu Gly Met Leu Ala Ser Pro Tyr His 145 150 155			741
TAC CTT ATC ATC ATA GTG GTT TTA GTC ATC ATT TTA GCT GTG GTT GTG Tyr Leu Ile Ile Ile Val Val Leu Val Ile Ile Leu Ala Val Val Val 160 165 170			789
GTT GGC TTT TCA TGT CGG AAG AAA TTC ATT TCT TAC CTC AAA GGC ATC Val Gly Phe Ser Cys Arg Lys Lys Phe Ile Ser Tyr Leu Lys Gly Ile 175 180 185			837
TGC TCA GGT GGT GGA GGA GGT CCC GAA CGT GTG CAC AGA GTC CTT TTC Cys Ser Gly Gly Gly Gly Gly Pro Glu Arg Val His Arg Val Leu Phe 190 195 200			885
CGG CGG CGT TCA TGT CCT TCA CGA GTT CCT GGG GCG GAG GAC AAT GCC Arg Arg Arg Ser Cys Pro Ser Arg Val Pro Gly Ala Glu Asp Asn Ala 205 210 215 220			933
CGC AAC GAG ACC CTG AGT AAC AGA TAC TTG CAG CCC ACC CAG GTC TCT Arg Asn Glu Thr Leu Ser Asn Arg Tyr Leu Gln Pro Thr Gln Val Ser 225 230 235			981
GAG CAG GAA ATC CAA GGT CAG GAG CTG GCA GAG CTA ACA GGT GTG ACT Glu Gln Glu Ile Gln Gly Gln Glu Leu Ala Glu Leu Thr Gly Val Thr 240 245 250			1029
GTA GAG TCG CCA GAG GAG CCA CAG CGT CTG CTG GAA CAG GCA GAA GCT Val Glu Ser Pro Glu Glu Pro Gln Arg Leu Leu Glu Gln Ala Glu Ala 255 260 265			1077
GAA GGG TGT CAG AGG AGG AGG CTG CTG GTT CCA GTG AAT GAC GCT GAC Glu Gly Cys Gln Arg Arg Arg Leu Leu Val Pro Val Asn Asp Ala Asp 270 275 280			1125
TCC GCT GAC ATC AGC ACC TTG CTG GAT GCC TCG GCA ACA CTG GAA GAA Ser Ala Asp Ile Ser Thr Leu Leu Asp Ala Ser Ala Thr Leu Glu Glu 285 290 295 300			1173
GGA CAT GCA AAG GAA ACA ATT CAG GAC CAA CTG GTG GGC TCC GAA AAG Gly His Ala Lys Glu Thr Ile Gln Asp Gln Leu Val Gly Ser Glu Lys 305 310 315			1221
CTC TTT TAT GAA GAA GAT GAG GCA GGC TCT GCT ACG TCC TGC CTG			1266

Leu Phe Tyr Glu Glu Asp Glu Ala Gly Ser Ala Thr Ser Cys Leu
 320 325 330

TGAAAGAATC TCTTCAGGAA ACCAGAGCTT CCCTCATTTA CQTTTTCTCC TACAAAGGGA	1326
AGCAGCCTGG AAGAAACAGT CCAGTACTTG ACCCATGCCC CAACAACTC TACTATCCAA	1386
TATGGGGCAG CTTACCAATG GTCCTAGAAC TTTGTAAACG CACTTGGAGT AATTTTTATG	1446
AAATACTGCG TGTGATAAGC AAACGGGAGA AATTTATATC AGATTCTTGG CTGCATAGTT	1506
ATACGATTGT GTATTAAGGG TCGTTTTAGG CCACATGCGG TGGCTCATGC CTGTAATCCC	1566
AGCACTTTGA TAGGCTGAGG CAGGTGGATT GCTTTGAGCT CGGGAGTTTG AGACCAGCCT	1626
CATCAACACA GTGAAACTCC ATCTCAATTT AAAAAGAAAA AAAAGTGGTT TTAGGATGTC	1686
ATTCTTTGCA GTTCTTCATC ATGAGACAAG TCTTTTTTTC TGCTTCTTAT ATTGCAAGCT	1746
CCATCTCTAC TGGTGTGTGC ATTTAATGAC ATCTAACTAC AGATGCCGCA CAGCCACAAT	1806
GCTTTGCCTT ATAGTTTTTT AACTTTAGAA CGGGATTATC TTGTTATTAC CTGTATTTTC	1866
AGTTTCGGAT ATTTTGGACT TAATGATGAG ATTATCAAGA CGTAGCCCTA TGCTAAGTCA	1926
TGAGCATATG GACTTACGAG GGTTGACTT AGAGTTTGA GCTTTAAGAT AGGATTATTG	1986
GGGCTTACCC CCACCTTAAT TAGAGAAACA TTTATATTGC TTACTACTGT AGGCTGTACA	2046
TCTCTTTTCC GATTTTGTGTA TAATGATGTA AACATGGAAA AACTTTAGGA AATGCACTTA	2106
TTAGGCTGTT TACATGGGTT GCCTGGATAC AAATCAGCAG TCAAAAATGA CTAAAAATAT	2166
AACTAGTGAC GGAGGGAGAA ATCCTCCCTC TGTGGGAGGC ACTTACTGCA TTCCAGTTCT	2226
CCCTCCTGCG CCCTGAGACT GGACCAGGGT TTGATGGCTG GCAGCTTCTC AAGGGGCAGC	2286
TTGTCTTACT TGTTAATTTT AGAGGTATAT AGCCATATTT ATTTATAAAT AAATATTTAT	2346
TTATTTATTT ATAAGTAGAT GTTTACATAT GCCCAGGATT TTGAAGAGCC TGGTATCTTT	2406
GGGAAGCCAT GTGTCTGGTT TGTGCTGCTG GGACAGTCAT GGGACTGCAT CTTCCGACTT	2466
GTCCACAGCA GATGAGGACA GTGAGAATTA AGTTAGATCC GAGACTGCGA AGAGCTTCTC	2526
TTTCAAGCGC CATTACAGTT GAACGTTAGT GAATCTTGAG CCTCATTTGG GCTCAGGGCA	2586
GAGCAGGTGT TTATCTGCCC CGGCATCTGC CATGGCATCA AGAGGGAAGA GTGGACGGTG	2646
CTTGGAATG GTGTGAAATG GTTGCCGACT CAGGCATGGA TGGGCCCCCTC TCGCTTCTGG	2706
TGGTCTGTGA ACTGAGTCCC TGGGATGCCT TTAGGGCAG AGATTCCTGA GCTGCGTTTT	2766
AGGGTACAGA TTCCCTGTTT GAGGAGCTTG GCCCCTCTGT AAGCATCTGA CTCATCTCAG	2826

AGATATCAAT TCTTAAACAC TGTGACAACG GGATCTAAAA TGGCTGACAC ATTTGTCCTT 2886
 GTGTCACGTT CCATTATTTT ATTTAAAAAC CTCAGTAATC GTTTTAGCTT CTTTCCAGCA 2946
 AACTCTTCTC CACAGTAGCC CAGTCGTGGT AGGATAAATT ACGGATATAG TCATTCTAGG 3006
 GGTTCAGTC TTTTCCATCT CAAGGCATTG TGTGTTTTGT TCCGGGACTG GTTTGGCTGG 3066
 GACAAAGTTA GAAGTGCCTG AAGTTCGCAC ATTCAGATTG TTGTGTCCAT GGAGTTTtag 3126
 GAGGGGATGG CCTTCCCGGT CTTGCGCACTT CCATCCTCTC CCCACTTCCC ATCTGGCGTC 3186
 CCACACCTTG TCCCCCTGCA CTTCTGGATG ACCAGGGTGC TGCTGCCTCC TAGTCTTTGC 3246
 CTTTGCTGGG CTTTCTGTGC AGGAGACTTG GTCTCAAAGC TCAGAGAGAG CCAGTCCGGT 3306
 CCCAGCTCCT TTGTCCCTTC CTCAGAGGCC TTCCTTGAAG ATGCATCTAG ACTACCAGCC 3366
 TTATCAGTGT TTAAGCTTAT TCCTTTAACA TAAGCTTCCT GACAACATGA AATTGTTGGG 3426
 GTTTTTTGGC GTTTGTTGAT TTGTTTAGGT TTTGCTTTAT ACCCGGGCCA AATAGCACAT 3486
 AACACCTGGT TATATATGAA ATACTCATAT GTTTATGACC AAAATAAATA TGAAACCTCA 3546
 AAAAAAAAAA AAAAAAAAAA 3566

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 386 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Leu Trp Gly Gln Ser Val Pro Thr Ala Ser Ser Ala Arg Ala
 -55 -50 -45 -40
 Gly Arg Tyr Pro Gly Ala Arg Thr Ala Ser Gly Thr Arg Pro Trp Leu
 -35 -30 -25
 Leu Asp Pro Lys Ile Leu Lys Phe Val Val Phe Ile Val Ala Val Leu
 -20 -15 -10
 Leu Pro Val Arg Val Asp Ser Ala Thr Ile Pro Arg Gln Asp Glu Val
 -5 1 5
 Pro Gln Gln Thr Val Ala Pro Gln Gln Gln Arg Arg Ser Leu Lys Glu
 10 15 20 25
 Glu Glu Cys Pro Ala Gly Ser His Arg Ser Glu Tyr Thr Gly Ala Cys
 30 35 40

50050936-053797

Asn Pro Cys Thr Glu Gly Val Asp Tyr Thr Ile Ala Ser Asn Asn Leu
 45 50 55
 Pro Ser Cys Leu Leu Cys Thr Val Cys Lys Ser Gly Gln Thr Asn Lys
 60 65 70
 Ser Ser Cys Thr Thr Thr Arg Asp Thr Val Cys Gln Cys Glu Lys Gly
 75 80 85
 Ser Phe Gln Asp Lys Asn Ser Pro Glu Met Cys Arg Thr Cys Arg Thr
 90 95 100 105
 Gly Cys Pro Arg Gly Met Val Lys Val Ser Asn Cys Thr Pro Arg Ser
 110 115 120
 Asp Ile Lys Cys Lys Asn Glu Ser Ala Ala Ser Ser Thr Gly Lys Thr
 125 130 135
 Pro Ala Ala Glu Glu Thr Val Thr Thr Ile Leu Gly Met Leu Ala Ser
 140 145 150
 Pro Tyr His Tyr Leu Ile Ile Ile Val Val Leu Val Ile Ile Leu Ala
 155 160 165
 Val Val Val Val Gly Phe Ser Cys Arg Lys Lys Phe Ile Ser Tyr Leu
 170 175 180 185
 Lys Gly Ile Cys Ser Gly Gly Gly Gly Gly Pro Glu Arg Val His Arg
 190 195 200
 Val Leu Phe Arg Arg Arg Ser Cys Pro Ser Arg Val Pro Gly Ala Glu
 205 210 215
 Asp Asn Ala Arg Asn Glu Thr Leu Ser Asn Arg Tyr Leu Gln Pro Thr
 220 225 230
 Gln Val Ser Glu Gln Glu Ile Gln Gly Gln Glu Leu Ala Glu Leu Thr
 235 240 245
 Gly Val Thr Val Glu Ser Pro Glu Glu Pro Gln Arg Leu Leu Glu Gln
 250 255 260 265
 Ala Glu Ala Glu Gly Cys Gln Arg Arg Arg Leu Leu Val Pro Val Asn
 270 275 280
 Asp Ala Asp Ser Ala Asp Ile Ser Thr Leu Leu Asp Ala Ser Ala Thr
 285 290 295
 Leu Glu Glu Gly His Ala Lys Glu Thr Ile Gln Asp Gln Leu Val Gly
 300 305 310
 Ser Glu Lys Leu Phe Tyr Glu Glu Asp Glu Ala Gly Ser Ala Thr Ser
 315 320 325
 Cys Leu

60050936.053097

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 331 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Leu Gly Ile Trp Thr Leu Leu Pro Leu Val Leu Thr Ser Val Ala
1           5           10           15

Arg Leu Ser Ser Lys Ser Val Asn Ala Gln Val Thr Asp Ile Asn Ser
          20           25           30

Lys Gly Leu Glu Leu Arg Lys Thr Val Thr Val Glu Thr Gln Asn Leu
          35           40           45

Glu Gly Leu His His Asp Gly Gln Phe Cys His Pro Cys Pro Pro Gly
          50           55           60

Glu Arg Lys Ala Arg Asp Cys Thr Val Asn Gly Asp Glu Pro Asp Cys
          65           70           75           80

Val Pro Cys Gln Glu Gly Lys Glu Tyr Thr Asp Lys Ala His Phe Ser
          85           90           95

Ser Lys Cys Arg Arg Cys Arg Leu Cys Asp Glu Gly His Gly Leu Glu
          100          105          110

Val Glu Ile Asn Cys Thr Arg Thr Gln Asn Thr Lys Cys Arg Cys Lys
          115          120          125

Pro Asn Phe Phe Cys Asn Ser Thr Val Cys Glu His Cys Asp Pro Cys
          130          135          140

Thr Lys Cys Glu His Gly Ile Ile Lys Glu Cys Thr Leu Thr Ser Asn
          145          150          155          160

Thr Lys Cys Lys Glu Glu Gly Ser Arg Ser Asn Gly Trp Leu Cys Leu
          165          170          175

Leu Leu Leu Pro Ile Pro Leu Ile Val Trp Val Lys Arg Lys Glu Val
          180          185          190

Gln Lys Thr Cys Arg Lys His Arg Lys Glu Asn Gln Gly Ser His Glu
          195          200          205

```

Ser Pro Thr Leu Asn Pro Glu Thr Val Ala Ile Asn Leu Ser Asp Val
 210 215 220
 Asp Leu Ser Lys Tyr Ile Thr Thr Ile Ala Gly Val Met Thr Leu Ser
 225 230 235 240
 Gln Val Lys Gly Phe Val Arg Lys Asn Gly Val Asn Glu Ala Lys Ile
 245 250 255
 Asp Glu Ile Lys Asn Asp Asn Val Gln Asp Thr Ala Glu Gln Lys Val
 260 265 270
 Gln Leu Leu Arg Asn Trp His Gln Leu His Gly Lys Lys Glu Ala Tyr
 275 280 285
 Asp Thr Leu Ile Lys Asp Leu Lys Lys Ala Asn Leu Cys Thr Leu Ala
 290 295 300
 Glu Lys Ile Thr Ile Ile Leu Lys Asp Ile Thr Ser Asp Ser Glu Asn
 305 310 315 320
 Ser Asn Phe Arg Asn Glu Ile Gln Ser Leu Val
 325 330

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 427 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Ala Gly Ala Thr Gly Arg Ala Met Asp Gly Pro Arg Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Leu Gly Val Ser Leu Gly Gly Ala Lys Glu Ala Cys
 20 25 30
 Pro Thr Gly Leu Tyr Thr His Ser Gly Glu Cys Cys Lys Ala Cys Asn
 35 40 45
 Leu Gly Glu Gly Val Ala Gln Pro Cys Gly Ala Asn Gln Thr Val Cys
 50 55 60
 Glu Pro Cys Leu Asp Ser Val Thr Phe Ser Asp Val Val Ser Ala Thr
 65 70 75 80
 Glu Pro Cys Lys Pro Cys Thr Glu Cys Val Gly Leu Gln Ser Met Ser

85	90	95
Ala Pro Cys Val Glu Ala Asp Asp Ala Val Cys Arg Cys Ala Tyr Gly		
100	105	110
Tyr Tyr Gln Asp Glu Thr Thr Gly Arg Cys Glu Ala Cys Arg Val Cys		
115	120	125
Glu Ala Gly Ser Gly Leu Val Phe Ser Cys Gln Asp Lys Gln Asn Thr		
130	135	140
Val Cys Glu Glu Cys Pro Asp Gly Thr Tyr Ser Asp Glu Ala Asn His		
145	150	155
Val Asp Pro Cys Leu Pro Cys Thr Val Cys Glu Asp Thr Glu Arg Gln		
165	170	175
Leu Arg Glu Cys Thr Arg Trp Ala Asp Ala Glu Cys Glu Glu Ile Pro		
180	185	190
Gly Arg Trp Ile Thr Arg Ser Thr Pro Pro Glu Gly Ser Asp Ser Thr		
195	200	205
Ala Pro Ser Thr Gln Glu Pro Glu Ala Pro Pro Glu Gln Asp Leu Ile		
210	215	220
Ala Ser Thr Val Ala Gly Val Val Thr Thr Val Met Gly Ser Ser Gln		
225	230	235
Pro Val Val Thr Arg Gly Thr Thr Asp Asn Leu Ile Pro Val Tyr Cys		
245	250	255
Ser Ile Leu Ala Ala Val Val Val Gly Leu Val Ala Tyr Ile Ala Phe		
260	265	270
Lys Arg Trp Asn Ser Cys Lys Gln Asn Lys Gln Gly Ala Asn Ser Arg		
275	280	285
Pro Val Asn Gln Thr Pro Pro Pro Glu Gly Glu Lys Leu His Ser Asp		
290	295	300
Ser Gly Ile Ser Val Asp Ser Gln Ser Leu His Asp Gln Gln Pro His		
305	310	315
Thr Gln Thr Ala Ser Gly Gln Ala Leu Lys Gly Asp Gly Gly Leu Tyr		
325	330	335
Ser Ser Leu Pro Pro Ala Lys Arg Glu Glu Val Glu Lys Leu Leu Asn		
340	345	350
Gly Ser Ala Gly Asp Thr Trp Arg His Leu Ala Gly Glu Leu Gly Tyr		
355	360	365
Gln Pro Glu His Ile Asp Ser Phe Thr His Glu Ala Cys Pro Val Arg		
370	375	380

Ala Leu Leu Ala Ser Trp Ala Thr Gln Asp Ser Ala Thr Leu Asp Ala
385 390 395 400

Leu Leu Ala Ala Leu Arg Arg Ile Gln Arg Ala Asp Leu Val Glu Ser
405 410 415

Leu Cys Ser Glu Ser Thr Ala Thr Ser Pro Val
420 425

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Gly Leu Ser Thr Val Pro Asp Leu Leu Leu Pro Leu Val Leu Glu
1 5 10 15

Leu Leu Val Gly Ile Tyr Pro Ser Gly Val Ile Gly Leu Val Pro His
20 25 30

Leu Gly Asp Arg Glu Lys Arg Asp Ser Val Cys Pro Gln Gly Lys Tyr
35 40 45

Ile His Pro Asn Asn Ser Ile Cys Cys Thr Lys Cys His Lys Gly Thr
50 55 60

Tyr Leu Tyr Asn Asp Cys Pro Gly Pro Gly Gln Asp Thr Asp Cys Arg
65 70 75 80

Glu Cys Glu Ser Gly Ser Phe Thr Ala Ser Glu Asn His Leu Arg His
85 90 95

Cys Leu Ser Cys Ser Lys Cys Arg Lys Glu Met Gly Gln Val Glu Ile
100 105 110

Ser Ser Cys Thr Val Asp Arg Asp Thr Val Cys Gly Cys Arg Lys Asn
115 120 125

Gln Tyr Arg His Tyr Trp Ser Glu Asn Leu Phe Gln Cys Phe Asn Cys
130 135 140

Ser Leu Cys Leu Asn Gly Thr Val His Leu Ser Cys Gln Glu Lys Gln
145 150 155 160

Asn Thr Val Cys Thr Cys His Ala Gly Phe Phe Leu Arg Glu Asn Glu

165	170	175
Cys Val Ser Cys Ser Asn Cys Lys Lys Ser Leu Glu Cys Thr Lys Leu		
180	185	190
Cys Leu Pro Gln Ile Glu Asn Val Lys Gly Thr Glu Asp Ser Gly Thr		
195	200	205
Thr Val Leu Leu Pro Leu Val Ile Phe Phe Gly Leu Cys Leu Leu Ser		
210	215	220
Leu Leu Phe Ile Gly Leu Met Tyr Arg Tyr Gln Arg Trp Lys Ser Lys		
225	230	240
Leu Tyr Ser Ile Val Cys Gly Lys Ser Thr Pro Glu Lys Glu Gly Glu		
245	250	255
Leu Glu Gly Thr Thr Thr Lys Pro Leu Ala Pro Asn Pro Ser Phe Ser		
260	265	270
Pro Thr Pro Gly Phe Thr Pro Thr Leu Gly Phe Ser Pro Val Pro Ser		
275	280	285
Ser Thr Phe Thr Ser Ser Ser Thr Tyr Thr Pro Gly Asp Cys Pro Asn		
290	295	300
Phe Ala Ala Pro Arg Arg Glu Val Ala Pro Pro Tyr Gln Gly Ala Asp		
305	310	315
Pro Ile Leu Ala Thr Ala Leu Ala Ser Asp Pro Ile Pro Asn Pro Leu		
325	330	335
Gln Lys Trp Glu Asp Ser Ala His Lys Pro Gln Ser Leu Asp Thr Asp		
340	345	350
Asp Pro Ala Thr Leu Tyr Ala Val Val Glu Asn Val Pro Pro Leu Arg		
355	360	365
Trp Lys Glu Phe Val Arg Arg Leu Gly Leu Ser Asp His Glu Ile Asp		
370	375	380
Arg Leu Glu Leu Gln Asn Gly Arg Cys Leu Arg Glu Ala Gln Tyr Ser		
385	390	395
Met Leu Ala Thr Trp Arg Arg Arg Thr Pro Arg Arg Glu Ala Thr Leu		
405	410	415
Glu Leu Leu Gly Arg Val Leu Arg Asp Met Asp Leu Leu Gly Cys Leu		
420	425	430
Glu Asp Ile Glu Glu Ala Leu Cys Gly Pro Ala Ala Leu Pro Pro Ala		
435	440	445
Pro Ser Leu Leu Arg		
450		

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 467 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Ala Pro Pro Pro Ala Arg Val His Leu Gly Ala Phe Leu Ala Val
1           5           10           15

Thr Pro Asn Pro Gly Ser Ala Ala Ser Gly Thr Glu Ala Ala Ala Ala
20          25          30

Thr Pro Ser Lys Val Trp Gly Ser Ser Ala Gly Arg Ile Glu Pro Arg
35          40          45

Gly Gly Gly Arg Gly Ala Leu Pro Thr Ser Met Gly Gln His Gly Pro
50          55          60

Ser Ala Arg Ala Arg Ala Gly Arg Ala Pro Gly Pro Arg Pro Ala Arg
65          70          75          80

Glu Ala Ser Pro Arg Leu Arg Val His Lys Thr Phe Lys Phe Val Val
85          90          95

Val Gly Val Leu Leu Gln Val Val Pro Ser Ser Ala Ala Thr Ile Lys
100         105         110

Leu His Asp Gln Ser Ile Gly Thr Gln Gln Trp Glu His Ser Pro Leu
115         120         125

Gly Glu Leu Cys Pro Pro Gly Ser His Arg Ser Glu Arg Pro Gly Ala
130         135         140

Cys Asn Arg Cys Thr Glu Gly Val Gly Tyr Thr Asn Ala Ser Asn Asn
145         150         155         160

Leu Phe Ala Cys Leu Pro Cys Thr Ala Cys Lys Ser Asp Glu Glu Glu
165         170         175

Arg Ser Pro Cys Thr Thr Thr Arg Asn Thr Ala Cys Gln Cys Lys Pro
180         185         190

Gly Thr Phe Arg Asn Asp Asn Ser Ala Glu Met Cys Arg Lys Cys Ser
195         200         205

Thr Gly Cys Pro Arg Gly Met Val Lys Val Lys Asp Cys Thr Pro Trp

```

210	215	220
Ser Asp Ile Glu Cys Val His Lys Glu Ser Gly Asn Gly His Asn Ile 225 230 235 240		
Trp Val Ile Leu Val Val Thr Leu Val Val Pro Leu Leu Leu Val Ala 245 250 255		
Val Leu Ile Val Cys Cys Cys Ile Gly Ser Gly Cys Gly Gly Asp Pro 260 265 270		
Lys Cys Met Asp Arg Val Cys Phe Trp Arg Leu Gly Leu Leu Arg Gly 275 280 285		
Pro Gly Ala Glu Asp Asn Ala His Asn Glu Ile Leu Ser Asn Ala Asp 290 295 300		
Ser Leu Ser Thr Phe Val Ser Glu Gln Gln Met Glu Ser Gln Glu Pro 305 310 315 320		
Ala Asp Leu Thr Gly Val Val Gln Ser Pro Gly Glu Ala Gln Cys Leu 325 330 335		
Leu Gly Pro Ala Glu Ala Glu Gly Ser Gln Arg Arg Arg Leu Leu Val 340 345 350		
Pro Ala Asn Gly Ala Asp Pro Thr Glu Thr Leu Met Leu Phe Phe Asp 355 360 365		
Lys Phe Ala Asn Ile Val Pro Phe Asp Ser Trp Asp Gln Leu Met Arg 370 375 380		
Gln Leu Asp Leu Thr Lys Asn Glu Ile Asp Val Val Arg Ala Gly Thr 385 390 395 400		
Ala Gly Pro Gly Asp Ala Leu Tyr Ala Met Leu Met Lys Trp Val Asn 405 410 415		
Lys Thr Gly Arg Asn Ala Ser Ile His Thr Leu Leu Asp Ala Leu Glu 420 425 430		
Arg Met Glu Glu Arg His Ala Lys Glu Lys Ile Gln Asp Leu Leu Val 435 440 445		
Asp Ser Gly Lys Phe Ile Tyr Leu Glu Asp Gly Thr Gly Ser Ala Val 450 455 460		
Ser Leu Glu 465		

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 343 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTCACGTTCC ATTATTTTAT TTAAAAACCT CAGTAATCGT TTTAGCTTCT TTCCAGCAAA	60
CTCTTCTCCA CAGTAGCCCA GTCGTGGTAG GATAAATTAC GGATATAGTC ATTCTAGGGG	120
TTTCAGTCTT TTCCATCTCA AGGCATTGTG TGTTTTGTTC CGGGACTGGT TTGGCTGGGA	180
CAAAGTTAGA ACTGCCTGAA GTTCGCACAT TCAGATTGTT GTGTCCATGG AGTTTTAGGA	240
GGGGATGGCC TTTCCGGTCT TCGCACTTCC ATCCTCTCCC ACTTCCATCT GGCCTCCACA	300
ACTTGTCCCC TGCACTTCTG GATGACACAG GGTGCTGCTG CCT	343

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 279 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTGGACGGTG CTTGGGAATG GTGTGAAATG GTTGCCGACT CAGGCATGGA TGGGCCCCCTC	60
TCGCTTCTGG TGGTCTGTGA ACTGAGTCCC TGGGATGCCT TTAGGGCAGA GATTCCTGAG	120
CTGCGTTTTA GGTACAGAT TCCCTGTTTG AGGAGCTTGG CCCCTCTGTA AGCGTCTGAC	180
TCATCTCAGA GATATCAATT CTAAACACT GTGACAACGG GATCTAAAAT GGCTGACACA	240
TTTGTCCCTTG TGTCACGTTT CATTATTTTA TTAAAAATT	279

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 250 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGCCACGTAG TGCCACGTGC CACAACTAC GGGGGACGAT TTCTGATTGA ATTTTGGCG	60
CTTTCAATCC ACCCTCCTCC CTTCTAATGG GACTTTGGGG ACAAAGGTCC GACCGCCTCG	120
AGCGTCGACA GGGCGCTATC CAGGAGCCAG GACAGCGTCG GGAACCAGAC CATGGCTCCT	180
GGACCCCAAG ATCCTTAAGT TCGTCGTCTT CATCGTCGGG TTCTCTGCCG GTAAGTTAGG	240
AGGTCCCTGG	250

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGCCCATGGC CACCATCCCC CGGCAG	26
------------------------------	----

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGCAAGCTTT TAGTAGTGAT AGGGAGAGGC	30
----------------------------------	----

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGCGGATCCG CCATCATGGG ACTTTGGGGA CAA

33

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGCGGTACCT TAGTAGTGAT AGGGAGAGGC

30

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGCTCTAGAT CAAGCGTAGT CTGGGACGTC GTATGGGTAG TAAGTGATAG GGAGAGGC

58